

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTAAGTACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTTGTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTTCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAACATGAAAGAAGCCATTTGATAGATTATCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIYSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTIILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGCCCCAGCGTGTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGG
TTTTCATGGTGCCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAAACAACGA
CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
CATTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTCTCAACTGCCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTACCACCTCAGCATCAGGAAGTGAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAACCTTCAAGGGGAAATATTTAATTTTCTTGGTT
ACTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTGTGATCGAGTT
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
GAAGTTTGGTCCCAACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCCTG
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
AGAATACCGCACCATAATCACTGAAGTCCTTGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTG
ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAATTTA
GATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGC
ATTTTTCACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVPPYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGBP
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALAEELSRQ
LFLETADLYATKERIEYSKTFKGKYNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGY
FVSSVLLIRMSPLEVRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

[illegible][illegible]

FIGURE 6

MARCFSLVLLLSIWTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQOKEMIETKVVKEEKANDSNPNEEKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGCGGGGAGCGAGCAGATCCAGTCCGCGCCGCGAGCGCAACTCGGTCCAGTCCGGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGTACCCGCGAGGAGGAGGCCACCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTATGAT
GGAGGACACGCGAGCACAATGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA
GCTGTGTGTCTGGGCTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA
CATCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTACAGCT
CCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGGGGGAGGAGATGGAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAAATG
CAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTCCACCTGCATTACATGTGTTTATTCATCCAGCAGTGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGTCTCT
CTCCACTACCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTCTTGGAGCATGCACATCTGGAATTAAGGTCAAATAATTCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTCATGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQMLCTRDSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQLEDLE
RSLTEEMALGEPAAAAAALLGGEET

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCACCTTGTGAACCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCTT
AACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAGATAGCCGGGTCATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAAGCCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCATAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTCTGCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGTCTGTCTTCTTTTTTTCTCCGGTCGCATCCCGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGGCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCCTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGTTTTAGGCCAGGCGCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGTCTCTATTAATAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAA
AAGATTTTATTAAAGATATTTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLEFWTL
NWVLALGQCVLGAFASFYWAFHKPDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVLDKVTDLLLFFGKLLVVGGVGLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCGGGCGCCCCGAAGCCGGGAGCCACCGCCATGGGGGCGCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCCTCATCTTCACGTTCTTCTCTTCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGGCTCCCTGGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCTTCTTCTTCTTCTTTTTCACCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCGCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGCTCGTGGG
CTCCTTCTCTTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTC
TTCTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACTGAGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGCCATCGCTGCTGCTCTGC
CCAAGGTCCAGGACGCCAGCCCACTCGGGTCTGCTGCAGGCCTCGGTATCACCTCTACACCATG
TTTGTCACTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG
TGGGCTCATCATCTTCTCTCTGTGCACCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGAGT
GGCAGCCTGTGAGGGCCGGGCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCC
ACTTCTGCCTGGTGTGCGCTCACTGCACGTATGATGACGCTCACCAACTGGTACAAGCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGTCCAGGACCTGCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGAACTCCCACCACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTTCTGGTCAAGTCCCCAGGGGACCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTCTAATAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCSCGSA PCILSCCPCASRNSTVSRLIFTFFFLGLVLSIIMLSPGVESQL
YKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFLILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV
QDAQPNSSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDPASI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGACCAGGAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGAAAAAGACTCCTGTAACCCCTCCTCCAGGATGAACACCTGCCAGAAGACATGGAGAACG
CTCTCACCAGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTAAAGTGTAAATACTTGCATATGCTGTGTGCAG
ACTGCCCATTTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTACTAGCAAAAG
TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTCATC
CTTGCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTAAATGTGAAAAACCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGCAGTGAGTCTCCCTGTGCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTTCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTATTTCATTCAC
AGAATGGAATTTTTTGTTCATGTCTCAGATTTATTTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTCTGTTTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCCCGCCGGCTCCCTGCGCCGCGCCGCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGCCAGGGCTGCCCAT
CCGGCTGCCAGTGCCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCCGGCCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACCAGATCGCCAGCC
TGCCAGCGGGGTCTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCGCCTGCTGCTGCTGGACCTCAGC
CACAAAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC
TGGATGTGTCCGACAACCACTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCCGGGCTGACG
CGCCTGCGGGCTGGCCCGGCAACACCCGATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
TGGTTTGGCCCTGGGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCGCCCAAGAAGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
CCACCACCACACAGCCACAGTGCCACCACGAGGCCGCTGGTGCGGGAGCCACAGCCTGTCT
TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC
CACTGCCCCACCGACTGTAGGGCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
ATGGGGGCACATGCCACCTGGGGACACGGCACCTGGCGTGCTTGTGCCCCGAAGGCTTCACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAG
GCCACCACGGTCCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGCTGGTGACGTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTCACCCA
GCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTCTATGCCTTTGGGGCCCGGGCGGGTGCCGGAGG
GCGAGGAGGCTGCGGGGAGGCCCATACACCCAGCCGTCCACTCCAACCACGCCCCAGTCACC
CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGAGCCTACTGTGTGCGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAAGTGGAGGGAGTGAAGGTCCCCTTGAG
CCAGGCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGCTGTAGTGTGAGGTGCC
ACTCATGGGCTTCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGACAGAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCAGTGCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT
GGGAAGATGTTTTCAAACCTCAGAGACAAGGACTTGGTTTTTGTAAAGACAAACGATGATATGAA
GGCCTTTTGTAAAGAAAAATAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLLLLLLLALGPGVQGCPSGCQCSQPQTTFCTARQGTTPRDVPPDTPVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDTANVE
ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNSLSLQALPGDLSGLFPRLRLAAARNPFNCVPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCEGFTGLYCESQMGQGRPSPTP
VTPRPPRSLTLGIEPVSPTSRLRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRGRAMAAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGLQSPHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGCCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAATGGAATGAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTCTATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAGTCAGGCAAAGGCTCTTGATATTATACAT
TTGGAGCTCTTGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTTTAGTGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTCAGCTTTCATGATC
CAGATTTGCTTGTTAAGACCAAATATTCAGTTGAACTTCCTTCAAATCTTGTTAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCTAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG
AAGATTAACTCATTTTTAATAAAATTATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDES LDSKTTLTSDSVKDHTTAGRVVAGQIFLDSESESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEPPKKVRKPALTAIEGTAHGEPCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMGEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAAGCCCATCTGCAGTGAATTTTCATGAACTAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTAATGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGT
CCCCGGCTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCGCTTGAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAGTCTCTCCCTAAGACTCATTATGCCGCTGAAAAGATGCCAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTTGCCCT
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCAGCATTACAGTAACCTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAA
AAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNOVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFCKTNLADPVKVIKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
CATTTCAGAAATATATTCATGCTATCAGCTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATCCGT
CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTCAGGAGCATT
TTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTTTTTACAGGGTACCTTTAGTGGTT
GCCAATCTGGGCATGTCTGAACAACTGGGTTATAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTTTCTTTGTGTCAGGCATTACGGACCTTTTTTCCAAATTCTGAATTTCTTCATTTCATGTGT
ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACCAACCACATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
CTACATTTTGATCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
ATTTCTATTGTTTTTACTATGTTGAGCTACTGTCAGTAAGTTCATTTGTTTTTACTATGTTACC
TGTTTGAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCTTTAGAATTGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
GAGACCATGTCTATTAAAAAATAAAATGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA
GAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLGGEVKGEAKNSITDSQMDDVEVVYTIDIQYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCCGCGCGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCAA
GCAGCGCGCAGCGAACGCCCGCCGCCACACCTCTGCGGTCCCGCGGCGCCTGCCACCCTTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGGAACCCCGAGGTACCCAGCCCGCGCCTCT
GCTTCCCTGGGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACCGGGGACCGTTGCTGA
CGCGAGGCCAGCTCTACTTTTCGCCCCGCTCTCCTCCGCTGCTCGCCTCTTCCACCAACTCCAACCTCTTCCCC
TCCAGTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAGCGCCGCTTCCCGTCCGCTCCAAA
GGTGGGAACGCGTCCGCCCGGCGCCGACCAATGGCACGGTTCGGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTC
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGTCTTTACGTGTCAAAGGCTTC
AACAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATCTTCAAAGAACTACTTGAAATGCAGAGAAATCCCTG
AATGATATGTTTGTGAAGACATATGCCATTTATACATGCAAAATCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGGAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCGCCCTGGTGAACCTCCAGTACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCCTCGCAAATGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCGCTACTTTGCTCAAGGC
TTAGCGGTTGCGGGAGATGTGCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGCTCTCGTGACTGTGAAGCCATGTTACAACCTACTGCTCAAACATCATGAGAGGC
TGTTTGCCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTTCAACATTGAATCGGTGATGGATCCCATGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGCAGCTGGCACT
AGTTTGACCGACTGGTTACTGATGTCAAGGAGAACTGAACAGGCCAAGAAATCTGGTCCCTCCCTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGACGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAACCCAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGCTCTGGGGCA
CAGGCCTACCTCCTCACTGCTTCTGCATCTTGTTCTGCTGATGAGAGAGTGGAGATTAATTCTCAAACCTCTGAG
AAAAAGTGTTCATCAAAAAGTTAAAGGCCACCAAGTTATCACTTTTCTACCATCCTAGTGAATTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGTGACTTTGTTTTCTCATTAGTTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTTCCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGATTGTGATTTTACTCTATTATTGTTTGTATGTTTTTCTCATTTCGTTTGTGGGTT
TTTTTTTCCAACGTGTGATCTCGCCTTGTCTTACAAGCAACCCAGGGTCCCTTCTTGGCACGTAAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTTATTAAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEELKYSLOSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAAKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCGTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGTGTTGTTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISL L L L L L P I M L M S M V S S S L N P G V A R G H R D R G Q A S R R W L Q E G G Q E C E C K D W F L R A P R R K F M
T V S G L P K K Q C P C D H F K G N V K K T R H Q R H H R K P N K H S R A C Q Q F L K Q C Q L R S F A L P L

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG
CTCTGTGCAGAGCCCAAGGAAGTCTGTCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC
GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGTTTTGTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLP LHL MALLGCWQPLCKSYFPYLM AVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTC LDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLC SVQSPRKVLQEVRRVLRPGGV LFFWEHVAEPYGSWAFMWQQVF EPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWL PVGPHIMGKAVKQSFPS SKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTCATTAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAAATCATGTCGG
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTGGCCATGATGTTTACC
TTCAGATTCATCACCACCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAAATGAAGTGCCTGTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTTCTGTCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
TTTTCTTCTGGGTCCCTCTGGGTGGCTGTGTCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGGAATATAAGCCCTTTCGGGCATTCCGTACATGTGGTGTACCATTTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCCTGCCAGCAAATGACTATAGCTGGGGCAG
TGTTTACTTGTATTTC AACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTCTGTCTCTC
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAAATCTCTGTGGTGAG
GATTCCGAGAATCATTGTCTATGTACATGAAAACGCACTGAAAGAACAGCAGCATGGTGCAATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAACCTGCTATTAAATGGGACAGATTTCTGTACATCAGCAAAGATGC
ATTCAAATCTTGTCCAAGAAGTCAAGTCACTTTACATCTATTAAGTCTTGGAGACTTCATAA
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAAGTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTT
TTAAAGACCTAATAAACCTATTCTTCTCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTDLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAQVWAVPLLLVAFFAYLVAHSFSLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGACCTCTCCCTGTTTCTTCCTTAGA
ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
GAAATTTTTTTGAAAAAAAATTCGCTTCTCAAACAAGGGTGTCAATCTGATATTTATGAGGAC
TGTGTTCTCCTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTGCTGGTGAAGTAC
ATTCAAACAAAGAAACGGCAAAGAAGATTAAAGGCCCAAGTTCACGTGTCCTCAGATCAACTGC
GATGTCAAAGCCGGAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACTACCCATCAGCTCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCAC
CACCTTGCCAAGGCCATCCCCCTTCTGCTGCTTCTACCACAGCATCCCAGACCACATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACTACACAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTGGAGC
GGATGTCAGCCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC
TGGGAGATCCAAATGCAAAATTGACTTGTGCTTTTTAATTGATGGGAGCACCAGCATTTGGCAA
CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCCAAGCTCTTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCTGCTACTCACTTTAACCTCAAGACAC
ACAGGAATTCCTGAGATCTGAAGACAGCCATAGAGAAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTCGGGCCATCTCCTTTGTGACCAAGAAGTCTTTTTCCAAAGCCATGGAAACAGAAGCGG
GGCTCCCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
GACTTGCGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAATGAG
AAGCAGTATGTGGTGGAGCCCACTTTGCAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
GCTCCAGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGTCTGCG
ACACTGACCGCCTGGCCTGCAGCAAGACCTGCTTGAACCTGGCTGACATTGGCTTCGTATCGAC
GGCTCCAGCAGTGTGGGGACGGGCAACTTCCGCACCGTCCAGTTTGTGACCAACCTCACCAA
AGAGTTTGAGATTTCCGACACGGACACGGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTTGGGTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCAGCGGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCCCTGGCT
GCCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTTCTTTGTGGACGA
GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
AGCCTCGGAACGAATTAGAGCAGGCAGAGCACCAGCAAGTGCTGCTTTACTAACTGACGTGTT
GGACCACCCACCGCTTAATGGGGCAGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
AAATGCTTGTATTATTCTTTGCCATCATGCTTTTTCATATTTCCAAAACCTTGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT
TTTGACAATTGTTTTCAAAATAAATGTTTCGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
AGCTTTTGTGAGATTTTAAAGTTGTTATTTCTGATTTGAACCTCTGTAACCCTCAGCAAGTTTCAT
TTTTGTGATGACAATGTAGGAATTGCTGAATTAATGTTTGAAGGATGAAAAATAAAAAAAA
AA
AAAG

FIGURE 34

MRTVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTPSALTYSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVS LGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQLPLVKRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLQLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAAQEELEVIATHPARHDSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATCCACAGTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCCTGGCTGCTCTATCACGTGGTGTCTCCGACTACTCACCCGAGTGTA
AAGAACCCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGAATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCAGTCTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTCTGGTCATTTC
TGGTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
TGTTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAGACAAAATGTTGGC
ATTGTCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTGAAACCATTATGGCATTGAGGTGGGTAAGTGAAGTTTGGCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATATAATGTCCAGAGAT
TTGGTGCCCAAGGATCTATGAAATGATGGGTACGTAAGCAATCCATCAAGTTTGAAGATGTTTATGTCCGGAT
CTGTTTGAATTTATTAAGTGAACATTCAATTCAGAGACACAAATCTTTCTTCTATATAGAATCC
ATTTGGATGTCTGTCACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCAAGGAGATCATCACTTTT
TGGCAGGTGATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGTGAAAGTGTTAAATAAGTAGGTACTGTGGAATTCATGGGGAGGTCAGTGTGCTGGCTT
ACACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAATATAAGGAATGGAGGTTTTTGCTAAAGAAATTAATAGG
ACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGTAAAAACAAAACATGTAGAGTTTATTTATTGAACAATGTAGTCACTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTTACCTGTTTTTGGTCATTATATAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT
ATTATTTAAATTAATCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAG
TGAATCATTCTTTACATGCAACATTTCCAGTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGACTTTGTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLLSLFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLT LKTIMAFRWVTEFCPNAKYVMKTD TDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQYEPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCCTGGAGCAGAGTGAATATCATGTGTCGGGAGTGTGTG
AATATGATCAGATTGAGTGCCTGCCCCGGAAGAGGGAAGTCGTGGGTTATACCATCCCTTGTGTCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAACGCAAGAGCTGCCGAAATGGCTCATGGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTCATGC
TAAACCTGGGTTTGTATCCAACTAAGATTGTTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTTCGTGATGGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTTCCATGACGGCAGTGCGTCTTGAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCGAGCCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCCCTGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAATTTGGCACCCTGGGTGCTTTCTTTTGAACACTCCTATGT
TCTTAGTGGCAATGAGAAAAAGAACTTGCCAGCAGAAATGGAGAGTGGTCAGGGAAACAGCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTAGACCTGGTGAGAAGGAGAGTTCCTCCGATGCAGGTTTCAAGGAGACACCATACACAGCTAT
ACTCAGCGGCTTTCAGAACGAGAACTGCAGAGTGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCCTTCTACCGCCGCTGGGCGAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGGCGGGCACCATCTGCATCCCTATCTGCGGGAAATGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGGGAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTTCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCTGAAGCT
CCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAGTCGGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAAGCAGACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCATGACAGACAGGAGGCATCGCGGCTGTGCTTCCCGGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGCCTGAAGTGTGATTGCGCTGTGAACCTGGCT
GTGCCAGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGTCTTCCCAACTTTCASTTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCTTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCCCTTCCCATCTCTTGTACACATTTTAATAAAATAAGGGTTGGCTTCT
GAACTACAAAAA
AAAAA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGDCMRGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGIENITAP
KTQGLRWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFAQESHITVAGWNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLPFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTGTATTTCAGAACTCTGTA
AAGTGCCCTCCACTACGGCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA
ACCTTGCTTACGTGTCTCGGCAGAGGACGGGAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAATA
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTAAAAAATAAATCGAGCTTTGAGTGTTCTTCGAA
GGACAAGAGCGGGAGTGCAATTGCCAACCATGCCGACAGGGCAGGGAATTTCTGAAAACACCACTGCCCTGAAG
TCTTTCGAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCAGTGAAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTATCGTGATGGGG
TGATGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTAGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCGCA
GCAGGAACATGGACAGGCCCCGATGCCACAGACCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTCATCTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCCAGAAAGTGGGCTCATCTGATTAGGCCAGTGAAAGACGTGTTACCTCGTGTGTCGCCAGGTTCCGCGAGC
GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCCAGGGGAGAGGCAACA
CTCCCAAGCCCTCCATCTACATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG
GCATGACCGTCGAGGGGAGCATCACATAGAGAATGGGATTGCTATCTATGTCATCAGTGTGAGCCCGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGTCGAACTGACAGAGGTCAGCC
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACACATGGCCCCACCCAGTGACTGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCACGGTGCTTGTATACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAATC
TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATAACAATGGAAACAAACCTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGCTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTCTTGGCCTGGCACTT
TTTTATGAATCAATGATGGGTGAGGAGAAACAGAAAAATCACAATAGGCTAAGAAGTTGAAACACTATATTTATC
TTGTGAGTTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAAAAATTAAGTACTAGTTTTTTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTTATATTTTTCTATTCAATAAAAGCCCTAAACAACTAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
AAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAACAGAATAAATATTTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSEGRSNRTRARPFFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSSENTAPEVFPRLYHLIPDGEITSIKINRVPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIENVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQSPDIFQEAGWNSNGSWSPG
PGRSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACIARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
TTCCACCTTTCTTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTCATGGCTAATTTCC
ATAAGACCCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACGTCTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
CTTTACAGAGGGTCCGCATCCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCTGCAGAGGCAGCAGTGGATTATGGCATCTACGTATCCACCAGGTGAAGG
TAAAAAGTTTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTACGTTACAGTGGATATTTGG
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTCAGGTGAATGGATTCTCTAACAACTACTGGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTCTGGTTTGGTGCATGACCC
TGATCTTTTGGTGATGTTTGGGAAGAACTGATTCTTTGTTTGCAATAATTTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCTTTTGTATTTTCT
TAGCAGAGCTCCTGGTGATGTAGAGTATAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAGGATGACTCAAAGGATAA
AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTATTTAAATTTGAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGT
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGTGCCCACCAGAGAATACATTCTCTATTAGT
TTTTAAAGAGTTTTTGTAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
ATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGENLTFHLSYKFRLLLLLTCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDCFIHFDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGVVTALSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCACTGGCTCTGTTTTCCCAACAG
ACGGGACAACCTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCCGTCCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTGTGTCACCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAAGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTGAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGACATTTGTGCGG
AACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAATCCCCACTCACTGCTCAGGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIIVTQNRNRERVDFFDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRILTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCAATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTTCATCCAGAATCCTTCAA
CTTGCACTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGAATTCCTGGAGGTCCTGTTGGGCT
CAGTCAGATAGTCATCGGTTTCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTS CNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCCGGCTCCTGGCAGGGG
GTGCTTGGCCACAGTGGTGTCTGGGAAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATAACCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCCAGCCTGGCTATGGTTCAGTGAG
AGCCAGCAACCAGAATGAAGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGGGGAGCGGGGAATCTGGGATTCAGGG
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
GCATCCCCGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACTCCCTCCTTAA
AACACCACCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
AA
AA

FIGURE 52

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVF
GHSGAWETSGGHGIFGSQGGGLGGQGNPGGLGTPWVHGYPCNSAGSFGMNPQGAPWGQGGNGGPPNF
GTNTQGAQAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSGGS
SSGSSSSSGSSGGSSGGSSGSGNSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
ARGSGESGIQGRGQGVSSNMREISKEGNRLLGGSGDNYRGQSSWGSGGGDAVGGVNTVNSETPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTGCCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAAGTCCGCCGGCTCCAGTGTTTCC
CACAGCCCCAAAACGGAAGTGGTTTTGGGGTCACTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCCTTCATCGTTTTATGCCACCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGACCCCAAGGATAATCTCTTCATCAGGTTCCCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTTGTAAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGCGTCGCACCTCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATGTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCAG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTCCGCA
GGGCCCAGGAAGTGCATCGGGCAGGCGTTGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT
GCTGCTGCACTTCCGTTCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCGAATGTAGGCTTGCACTGACTTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEKILLSGGDKWSRHRRLTPAFHFNIKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDVLDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG
TACAGCCTTTTCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTGCTTGC
ATAATTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTTGCTT
GTGATTCTGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGGGAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATCCCATGAATACAAACCTATTTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCGGCCAACATGCTCTGTCTGTGCCTG
TACGTGCCGGTTCATCGGGGAAGCCAGACCGAGTTCCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTCATCCCTCCAGGAATTTCCACCTACCGCCAGTGGAGCAGAAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCGAGTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTTTAAGATTTTGGACAAAAAAGATGATGGACGCATTGACGCGCAGGAGATCATGCACTCCCTGCCGGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCGAGAAAAATTTCTCAAGAGCATGGATAAAACGGCAGCATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCGCTGGAAAAACATCCCGAGATCATCCTCTACTGGAAGCATTCCACG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCTCCCGCAGCAACAATGCGCATCGTTGGTGGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGGCAATGGCATCAACGTCTCAAAATTTGCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACAGGAGACTCTGAGGATTACGAGAGGGCTTGTGGCAGGGTCTTGGCAGGGGCC
ATCGCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCGCCAGGAGGATCTGGCCAGAGAGGGGTGGCCGCTTCTACAAAGGCTATGTCCCAACATGTG
GGCATCATCCCTTATGCCGGCATCGACCTTGCACTACGAGACGCTCAAGAATGCCTGGCTGCAGCATTATGCACTG
AACAGCGCGGACCCCGGCTGTTTGTCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAAACATATCTGCGGACCGAGGGGGCTTCGGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGTTCATCCCA
GGTGTGAGCATCAGCTAGTGGTCTACGAGAACCTGAAGATCACCTGGCGGTGCACTGCGGGTGAAGGGGGAGGGC
CGCCCCGAGTGGACTCGCTGATCCTGGGCCGAGCCTGGGGGTGTGAGCCATCTCATTCTGTGAATGTGCCAACAT
AAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGAGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTT
GTCTGTGACCCAGCAGACCCTCCTGTTGGTTCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGG
CTCCGGGCTCACATGTGTAAAGACAGGACATTTCTGCACTGCTGCCAATAGTAGCTTGGAGCCTGGAGGCCGGCT
TAGTTCCTTCCATTTACCCCTTGACCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTTGTGCTGCTGCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAGCCACATCCCAACCCCTCGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAACTTGGC
TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAGTGCAGGGGGCTCGGGCTGCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT
GGCCTGGACCTGTGAGATGGGCCCCACCTCAGAACCACACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAAGGCGAAGGGCAGAGCGTTTGTGTCTTGGGGAGGGAAGGAAAGGTGTTGGAGGCCCTAATTATGG
ACTGTTGGGAAAGGGTTTTGTCCAGAAGGACAAGCCGGACAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTCTGTCCAACCCAGCAGGGGGCGCAGC
GGGACCAGCCCCACATTCACCTTGTGTCACTGCTTGGAACTATTATTTTGTATTATTGAACAGAGTTATGTCTT
AACTATTTTATAGATTTGTTTAATTAATAGCTTGTATTTCAAGTTCATTTTATTATATTTATGTTTATGTTT
GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGAGAAGGGGGGCTTGGGGCGCTGCAGTCACATCT
GTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAAGGCAGCAGCCTGGCTCCTTCTCTTGGCAG
GTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTAGGGTTTGAAGTGGGGCGTGGAGAGAGGGGAGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTATTTTCACTCTTTTCTGAATGTCAAGGCAG
TGAGGTGCTCTCACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGCTGGCTCCGCTCCCTCCAGC
CTTCTGCTGCCCTTGTCTTAAACATGCCGGCCAACTGGCGACCTCACGGTGTCACTTCCATTCACCAGAATGACCTGA
TGAGGAAATCTTCAATAGGATGCAAGATCAATGCAAAAATTTGTATATATGAACATATAACTGGAGTCGTCAAAAAG
CAAATTAAGAAAGAAATGGACGTTAGAAGTTGTATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQIVQAGDKDLDG
QLDFEETFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEI MQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRKVLVMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSQDET LRIHERLVAGSLAGAI AQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQA QASIEGAPEVTMSSLFKHILRTEGA FGGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTTGAACCTGACATCAAACTTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCTGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTTCGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCCTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAAATAAT
GTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC
CACCAGATATGACCTAGTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAAGTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCCTGGAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACCTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
TCTTCCCACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTESIARRSHLQLLNSKASL
CVSSFFAISWALLPLSPYMLLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTGGTTATGGGGGTCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGTCCCGGGGGCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAGAAGCCTCTT
CCAGCAGTATGACCGGGACCCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATGACCCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTCTTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAATGG
AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCCTGTACC
GTTAGTACCTGTGTCCCTCACCAGGCCATCCTGTCAAACGAGCCATTTTCTCCAAAGTGGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCATAAAT
CCTTGTTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTTGCATTTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAATTTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYPERSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCAAAACCTTCGGGACC
AACTTGCCTCAGCTCGGACAACCTTCCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCCG
TCTGGACCCCTAGGTCTAATGACTTGGCAAGGGTTCCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACTCTTTACCTCTCCAGTGTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
TGCTTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAACCCCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNLSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCCTTGTCTCTCTTGACCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGGCCC
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGPLPLVLLLTLLGSSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC
CAGGTGCCCCGTGCGAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCCTCGGGCGTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGGCTGCCGTTCTGCTGGCCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACCTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLRLARWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTAGAGGTTTGTCTTCTAGTCTGTGCCTGCTGCACCAG
 TCAAATACTTCTTCATTAAGCTGAATAAATGGCTTTGAAGATATTGTATTGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
 TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAACATGAA
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCAGGTAGAGATGAACCATACCAAGCAGTTCACA
 GAATGTGGAGAGAAAGCGAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA
 CAGGCCAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAATGTATGGAAAAAGATTGTCAATCTTT
 CCTGATAAGTACAAACAGAAAAAGCATCCATAATGTTTATGCAAAAGTATTGATTCTGTTGTAATTTTGAACGAA
 AAAACCCATAATCAAGAAAGCTCCAAGCCTACAAAACATAAAGTGAATTTTGAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCCTCTCCACCTGCTTCTCATTGCTGAAGATCAGTCAA
 AGAATTGTGTGCTTAGTCTTGTATAAGTCTGGAAGCATGGGGGTAAAGACCGCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGTGCTCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCATTTGATAGTACTGCCCATTTGTA
 AATAAGCTAATCCAAATAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
 ACTTCCATCTGCTCTGGAATTAATATGCATTTCAAGTGATGGAGAGCTACATTCCTCAACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTCTCTGTATGATGAAGTGAACAAAGTGGGGCCATTGTTTAT
 TTTATTTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAGG
 TCCCTTCAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCAATTTGATAGTACA
 GTGGGAAAGGACACGTTCTTCTCATCACATGGAACAGTCTGCCTCCAGTATTTCTCTCTGGGATCCAGTGGGAACA
 ATAATGGAAAAATTTACAGTGGATGCAACTTCCAAATGGCCTATCTCAGTATTCAGGAACGCAAGGTGGGCACT
 TGGGCATACATCTTCAAGCCAAAGCGAACCCAGAAACATTAACTATTACAGTAACCTTCTCGAGCAGCAAAATCTTCT
 GTGCCTCCAATCACAGTGAATGCTAAATGAATAAGGACGTAACAGTTCCTCCAGCCCAATGATTGTTTACGCAGAA
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCATTGAATCACAGAATGGACATACAGAAGTT
 TTGGAATTTTGGATAATGGTGACGGCGCTGATTCTTCAAGAATGATGGAGTCTACTCCAGGTATTTACAGCATAT
 ACAGAAAAATGGCAGATATAGCTTAAAGTTTGGGCTCATGGAGGACAAACACTGCCAGGCTAAAATACGGCCTCCA
 CTGAATAGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAATGAAGCAAACCCGCCAAGCCTGAAATTTGAT
 GAGGATACTCAGACCACCTTGGAGGATTTCAAGCGAACAGCATCCGGAGGTGCATTTGTGGTATCACAAAGTCCCAAGC
 CTTCCCTTGCTGACCAATACCCACCAAGTCAAATCACAGACCTTGATGCCAGGTTTATGAGGATAGGATTATCTT
 ACATGGACAGCACCAGGAGATAATTTTGTATGTTGGAAGTTCACAGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTACCAAAGGAGGCCAACTCCAAGGAA
 AGCTTTGCATTTAAACAGAAAAATATCTCAGAAAGAAATGCAACCCACATATTTATTTGCCATTAAAAATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACTTGCAAGTAACCTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGAT
 CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTTCTGGAGTTAATATTCTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATTTGTTAACTTTATTTTAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTT
 AAGTAGACCTAGAAGAGAGTTTTAAAAACAAAACAATGTAAGTAAAGGATATTTCTGAATCTTAAATTCATCCCAT
 GTGTGATCATAAACTCATAAAAAATAATTTTAAGATGTCGGAAGAGGATACTTTGATTAAATAAAACACTCATGGATA
 TGTAAAACTGTCAAGATTAATTTAATAGTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAACAAAG
 ATCCTTTTTCATCTGATACCTGGTTGATATTTATTTGATGCAACAGTTTTCTGAAATGATATTTCAAATTCATCAA
 GAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAATAAACCAACATTTGAAAAAATAAAAA
 AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLKGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRNLNR
MNQAAKHFLLOTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLLTGDGENTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTIITVTSR
AANSSVPPITVNAKMNDVNSFSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNDGVYSRYFTAYTENGYSKLVRAHGGANTARLKLRLPPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLEDFSRASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWAPGDN
FDVGKVQRYIIRISASILDRLDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGTGGAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGAGGGGTGA
CAACAGGTGTCACTTTTTGATCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTACTAGAGAGTGTCA
ACTCAACAGCAAGCGGCTCCTTCGGCTTAACCTTGCTTGGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCTCTTGTGTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAA
AGTAGAGAAGCTGCTCTGTGTGGTGTAACTCCAAGAGGCAGAATCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGATGATGTTTCGCCGGGGGCTGCTTGCCTGGATTTCCTGGGTGGT
GTTTGTGCTGGTCTCTCTGCTGTGCTATCTCTGCTGTACATGTTGGCTGCACCCCAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCCAACAGCCCCACGGGGAAGGAGGGGTACAGGCCGTCTTTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGG
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGACGAGTG
CCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAAAGCCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGGAGCCATTGAATCAGCCTTGGAGACCTGAAACAATCTGCAGAGAACAGC
CCCAATCACCGTCTTACAGGCCCTCTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGACAAAGGGACATTGTAT
GAGCTCACCTTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTGACCATTCAGCCCCATCATGAAA
GTGAAAAATGAAAGCTCAACATGGCCACACGCTTATCAATGTTATCGTGCTCTAGCAAAAAGGGTGGACAAGTTC
CGGCAGTTCATGCAAGATTTTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
AAAGAAGAAATAAATGAAGTCAAAGGAATCTTGAACACTTCCAAAGCTGCCAATTTCAGGAACTTTACCTTCATC
CAGCTGAATGGAGAATTTCTCGGGGAAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGCTCTCTC
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
GTATTTTATCCAGTTCTTTTCAGTCAGTACAATCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTTGGAA
CAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTGAGTATCGGTGAGAC
TTATCAATATAGGTGGGTTTGTCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCCGTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCGGAGCAGTACAAGATGTGCATGCACTCCAGGCCATGAACGAGGCATCCACCGCCAGCTGGGCATGCTG
GTGTTTCAGGCACGAGATAGAGGCTCACCTTCGCAAAACAGAAACAGAAAGCAAGTAGCAAAAAACATGAACCTCCCA
GAAGGATTGTGGGAGACACTTTTCTTTCTTTTGCATTAATCTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAGAATTGGACTGATGGGTGAGAGATGAGAAAAGCCTCCGATTTCTCTGTGTTGGGCTTTTACAAACAGA
AATCAAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTGACAAAGGCAGAAATGCTTGTG
AGATTATAAGCCTAATGTTGTGGAGGTTTTGATGGTGTTCATAATACATGAGACCTGTTGTTTTGTGTGCTCATTGA
AATATTCTAGTTTAAAGCAGTTTGTGAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAAATATCAGAAGGCAGGAGAGAGATAGGCTTATATGATACT
AGTGAGTACATTAAGTAAATAAAATGGACAGAAAAGAAAGAAACATAAATATCGTGTCAATATTTTCCCCAAGAT
TAACCAAAAAATAATCTGCTTATCTTTTGGTTGTCCTTTTAACTGTCTCCGTTTTTTCTTTTAAAAATGCACT
TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC
ATTTTTATATTTTTTAAAGATACTTTGAGATGCATTATGAGAACTTTCAGTTCAAAGCATCAAATGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCACTGAGACATAGGGAAGGAATGGTTGTACT
AATACAGACGTACAGATACTTTCTGGAAGATATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
ACTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAGTCAGAAACCACTTTT
CTCCTCAGAAGTAGGACCGCTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTTC
AAAACAGGGTCTCCTCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATATTTGT
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAGTTTTTGTCTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACTGAATTATTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTT
CAAACTATTTTTTAAAAATAAACAGTTAAACATAGAGTGGTTTTCTTATTGATGTGAAAATTATTAGCCAGCACCAG
ATGCATGAGCTAATTATCTCTTTGAGTCTTGGCTTCTGTTGCTCACAGTAAACTCATTTGTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTTGACTGGTAGTTTATGAAATTTAATTTAAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVNEKLNMAN
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLSQY
NPGIIYGHDAVPPLEQQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYR
KYLHSNLIIVRTPVRGLFHLWHEKRCMDELTPEQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGATCCA
GAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTCTGCCTGTCTAGTCTCTAGTCTCTCAAATCCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCCTGGCTGCAGATGGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGGCCAGCCTCTTACCCTGAGTGT
GGAAACAATGCCAGTCGCCCATCGATATTCAGACAGACAGTGTGACATTTGACCCTGATTTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCAGCCTTTGGACCTGCACAACAATGGCC
ACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTGGGTGGACTTCCCCGAAAATATGTAGCT
GCCCAGCTCCACCTGCACTGGGGTCAGAAAGGATCCCCAGGGGGGTGAGAACACCAGATCAACAG
TGAAGCCACATTTGCAGAGCTCCACATTGTACATTATGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCTCAGGGCCTGGCTGTCTGGCATCCTAATTGAGGTGGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGTCATGAAGTCAGGCATAAAGATCAGAAGACCTC
AGTGCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAACAGCTGGGGCAGTACTTCCGCTACAATG
GCTCGCTCACAACCTCCCCCTTGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAG
ATTTCAATGGAACAGCTGGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTGGC
TGTCTCTGCCTTCTCCTGGCTGTTTATTTTATTGCTAGAAAGATTGGAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCCTTCTCAGATAC
CATGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGGGAAGTTGGG
ATATACCCCAAAGTCCTCTACCCCTCACTTTTATGGCCCTTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTGTATCAATATATTTGGAAATTAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCGCTGCTGCTGTTGCTCCTGGCGGCCTTGGGGACGGGCAGTTCCTGTGTCT
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC
AGAGAAGTGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCCTTCGTCCC
AGGGCCCCCTCGCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
GATTTTGATTTATGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTACCCAGCAAGAGTCCCTCAGCAGACAATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTGAGAACCCTGACATTTGTGCGGGGCTGAAGAGCAGGA
GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCTGGTTCGACTGGGA
TCCCCAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTGAGAGGCTGCG
AGCCTTCTGAGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATA
TGTGCAGATGGAAACTGATGCCAAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAG
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT
CTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTTCATGCATGTA
GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTGTCTATGCAGAGAA
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYIGNEFDKRFFVPAEK
IVINFITLNISSDKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDKTVIEYDYVRTTDICAGPEEQELSLQEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGGCCGGCCCGTGGACCTTCACCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCCCTCAGTCCCACTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAAGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCAATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCACCCGCCTGGTCTCAGTGAAGTGTGCCACCAGCCATGGGAGCCTGGCATCCAAGTGTGTA
TAAGCTCTCCTTCTGGTGAAGCCTTAGCTAAGCAGGTCAAGACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAACAGCTGTGTCCCCTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTTCGAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTGATGGTCTGTTGGACTCTGTGCTTCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTGTTCACCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGGAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPKIGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGGTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGAGGCCTCCTGGGATTC
ATTCTGTTCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCAGTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCTGCTCATCCAGAGAAATCGCTCCAACTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGCAATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCTCAACTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCCTGACTGACCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTGTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAAGTGAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVG YILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWN LHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFN SYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC
CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTTCCT
GCCCACCGCTGCTTCTGGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCTCCCGACTCCGCTCCCGG
ACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCTCTCCTTCTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCCTGCCCAACCAGTGTGTCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAACTGCCCCGAACCAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGTGGGAGAAAG
AGAGGCCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAAGTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG
ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGAATCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTTCTAGCCCAGACCCTGGAGCTGAAGGTACAGGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRTCSE
GAHVSCYRLHCPVHCPQPVTEPQQCCPKVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTAATCCATCCGTCACCTCTCCTGTCA
TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGAG
GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGGCGTTCTTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
ACAGGTGTGAGCACTGGGCTCAGTTCCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCACTCCTCGGGCTGGTTCCTCCCGGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGTATGTGGAGATCTCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAATTCAGTGGAAAATCCAGGCGGAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAAATTGAGAGACGCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTTC
ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTCTGGACTA
TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTGCGTTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTTATAATGAGCAAAATGGAACCTCCCATAGTCTATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG
TGAGTCTCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTTCTTTAGGGATATTAAGGTCTCTCTCCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
GATGAAGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCTGAGTGGCTGAGCTGGGAGGGAAGAAGG
CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
AACCGTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGTTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MAIMLSLVLSLLKLGSGQWQVFGPDKPQALVGEDAAAFSCFLSPKTNAEAMEVRFFRGQFSSVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQSFOAGKHYWEVDGGHNKRWRVGVCRDDVDRKEYVTLSPDHGYWVLRNLNGEHLFT
LNPRFISVFPRTPPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPTSNSSESSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

[illegible]

Author	Year	Country	Sample Size	Study Design	Findings
Wong et al.	2001	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2002	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2003	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2004	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2005	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2006	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2007	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2008	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2009	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2010	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2011	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2012	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2013	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2014	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2015	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2016	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2017	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2018	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2019	China	1,000	Case-control	Increased risk of lung cancer with tobacco use
Wong et al.	2020	China	1,000	Case-control	Increased risk of lung cancer with tobacco use

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPDHGTS LTCQVTFPGASVTNKTVHLNVSYPQNLMTVTFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVS LQSKATSGVTQGVVGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGD TGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGGCGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTTAT
GGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLEFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIIPVYDFGDAQKTASYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCCATGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCCTTCTTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCACTCTGTGGCCAGGATGATG
GTCCTCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCTTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCGAACCACAGCCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCAGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACACTAC
CATAGTGATACCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGGCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCTTGTCGTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCCACAGCCCATC
CGCGTGCTGTGTGCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTT
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLLYLICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCTAT
GGAGATTCTCCTCAATTCTTCCTTTATTTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGTATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTCTAGCCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGTCTATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPGTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFILAAGGSCRSLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAAAC
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTGCACAACTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTTCCTTGTTCCCGGGAGGCATCTGCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCGCTCCCAACTCCCAGTGGCACAGATGACGACTTTCAGTGACACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAAAACCTGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGFDLHLLNPAAGMTPGTQTHPLTLGGLNVQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

1000
900
800
700
600
500
400
300
200
100
0
100
200
300
400
500
600
700
800
900
1000

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTGACCATG
GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCCAAGGCCAGCCTGC
AGAGCTGTCTGTGGAAGTTCCAGAAAACATGGTGGAATTTCCCTTTATACCTGACCAAGTTGC
CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTGAGGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGCTGGTGACCAGGGCCCTGGACCCGAGAGGA
GCAGGCAGAGTACCAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGGTACCAGGCCCTGGCATCCCCTTCTCTTCTTGGAGCTTCAAGCCG
GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCCAGGCACTGCCACCGTGGAGTCTCCATCATAGAGAGCAGCTGGGTG
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGACCCACATGGCCAGGTA
CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCCCTTTGAAGTGAA
TGCAGAGGGAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTCCTATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTG
ATGGATGAGAATGACAACGTGCCATCTGCCCCCTCGGCTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGACTAGACTGTGACAGAGGATGCAGATGCCCCGGCTCCCCCA
ATTCCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGGCCTTC
CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
CCTGCTTCTGGTGTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAAGCAGCAGTGTGAAG
TCGAAGTCGAGTACAGATATCAATGATCAGCCCCCTGAGTTCATCACTTCCCAGATTGGGCCT
ATAAGCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
CCTCGAGCCCGCCTTCCGCTCATGGATTTTGCCATTGAGAGGGGAGACAGAGGGGACTTTTG
GCCTGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCCTCAGTTATGAG
GCAGTCCAAGTCATGAGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCC
AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGTGCCACCCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTCTCTGCTGACCATC
CAGCCCTCCGACCCCATCAGCCGAACCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGGCACGGTCCCTACAGCTTACCCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC
GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCAGT
GAACACATAATCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGAT
CGTGTGCTGCTGCAACGTGGAGGGGAGTGCATGCGCAAGGTGGGCGCATGAAGGGCATGCCCA
CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTC
ATTTTACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
GAAGGGGACTGTCTGAATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
TCCCCCTGGGAGAGAGCCAGCACCCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
TCTGCCCTGGGGTGGAGGCCACCATCACCATCACAGGCATGTCTGCAGAGCTGGACACCAACTT
TATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTGGCCCAATAATAAAGCCCCAGAGAA
CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAQYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGNPTVQRDWRLQTLNGSHAYLTALHWWEP
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTILVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
TCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTCTTTGCCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
GCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCTCCAGTGGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACACCCTCC
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGTGAGGCAAAGCCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACACCCT
CATGGCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCACAGGCCAG
GTGGAGTCCTAAGTGGTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGGCGCATTCTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTTCTTTTATTTCATCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAA
AATCTTGAAGAAGGTATTCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACAGACAAAGAGAAGCTGTGCTTG
CCCCGGGTGGGTATCTAGCTCTGAGATGAATCAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG
GGAGGCTGGAACATACAGGCTGAAAAACAGAGTGGGTACTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTGATGGAAGGGATTACTTTATACTGACTCTGTTTTGGGAAGCTTTTTTGGAGCATT
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAGT
GATTATAACTGGGGATGCATTTGTCTCGGAGAAAGAAGTGTGATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTCGGATTTGGTTGGGCCATGCAGGCTGCTGCCATAT
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTTGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTTACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTGATTCCACCTTGCAAGTCTGAAGTCAAGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTCAAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTAAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCAATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTGCATGACTATGTCGAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGCACCTAATTTTGTGGGAAAAATATTGCTACAATTTTT
TTAATCTCTGAATGTAATTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFSGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGWAMQAAAYIFIHRKWDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLLQGDFPREIHFFVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVLVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIEELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTTGTGCTCATTGTGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
TTCCTGAGACCACTGCAGAGGCTATGGAAGTGCAGTTTCTCAGGAATCAGTTCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGAAAGGTCCACAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC
AAATCCAAAGGGAAAATCCAGGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTACCCGAAGCTCTGCG
TTTCTGATCTGAAACTGTAACCATAGAAAAGCTCCCCAGGAGGTGCCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA
ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT
ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGC
TGACATGTGAGTTTGAAGGCTTGTGAGACCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCCCGACAGGTGGCCCCAGCTTCCTCT
CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTGTGCTAGTCACGG
ACAGTGATTCTGCTCAGAGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTTATATTATACATTTTCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACCAGGCAAGAAAACAGAAGAAGAGGAAGGAAAACACAGGTCCATATCC
CTCATTAACACAGACACAAAAATCTAAATAAAATTTAACAAATTAACATAACAATATATTTA
AAGATGATATATACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTTAAATATTTAAAT
ATCAACCAGTGAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQSPWRLASILLGLLCCALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCAATTTACAAC
TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATGCTGCGGA
ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
TTGTGAGTGTCTCACTGTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCTTTGGAGTAACA
ATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAATACAAACACCCATCACA
TGACTATGATATTTCTCTTGAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTG
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACTTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGGTGACTCTGGAGGACCCTGGTTAGTTTCAAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTG
GATTACTTCAAAAATGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTGT
GGTGTGGAGGCCATTTTATAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAACTGTTTGCTTGATGCATGTATTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCTCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCAATTTGTTCTCTAGAAGTTTGTGAGAATTTTGACTTGTGACATAAATTTGTAAT
GCATATATACAATTTGAAGCACTCCTTTCTTCAGTTCCTCAGCTCCTCTCATTTTCAAGCAATATCCATTT
TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTCTTAGTGGAATATTAGAAATGATCATATTCAATATGAAAGGTCAAGCAAAGACA
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCT
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYNSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLOWDGSHRGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSFPYTNVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSSGGLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGGAAGAAGTTCCCTG
CCCCGATGAGCCCCCGCGTCCCGGACTATCCCAGGCGGGCGTGGGGCACCGGGCCCAGC
GCCGACGATCGCTGCCGTTTTGCCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
ACGGGGCTCACAAATGCGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCCT
GTGGTTCATCCGGTCATGATTGCTGTTTGCTGTTTCCTTATCATTGTGGGGATGTTAGGATATTG
TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCAATGTTGCTTGGAGTTGCTTGTCAATTTCT
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCAGTACAATGGTCA
GATATGGTCACTTTGAAAGCCAGGATGACAAATATGGATTACCTAGATATCGGGTGGCTTACTCA
TGCTTGGAAATTTTTTTCAGAGAGAGTTAAGTGTGTGGAGTAGTATATTTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCCGTGTGTGTAGAGAATCCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTGTGGGAAGAAAATGTATTCCTT
TTTGAGAGGAACCAACAACACTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCTGGG
ACAGACCAAATGATGCTCTTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
TTGAGATGGAGGAGTTATAAAAAGAAATGTCACAGAAGAAAACCAAACTTGTTTTATTGGACT
TGTGAATTTTGTAGTACATACTATGTGTTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
TAACACCTAAGCATATACTATTCTATGCTTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCGCCA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTTCTACCAA
CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
ACCTAAATGTGATTTTGTGTTACTTAAATATCTTACCCTTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCACTCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGCTCCTGTATAGCATCATT
ATTTTGTAGCCTTTCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
TGTTATTTAAATACTTAACCACTAATTTTGAATAATACAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
AAGGACTTGTATGCTGTTTTCTCCCAATGAAGACTCTTTTGGACACTAAACACTTTTTTAAAAA
GCTTATCTTTGCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTTACAGAAAA
TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATT
CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTACAA
GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCAATTTTACTGTATTTTGTGATTTTGTATT
TTCTCAGAAATATGAAAGAAAATTTAAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKMYSLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCTCTGATAAAGCCCTACCACTGCT
GATAAAGTCTTTCTCTGAGAGCCTAGAGGCCTTAAAAAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTTCCAGTGTCTTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTACCCGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTTGGATGAGTCTCAAGGTCTCCTCTATGTGACAACCATGTGAATGGGGAG
TGGTACCATTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGGAACGTGTCTCTGGAACACCACGGTGAAGTCAAGGCTTGCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAACGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCCTGT
GAGTGTGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCTTGGATCTGAGAAAGGCTACCAAGTGTGAATGTCCCGGGGCTTGGTCTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCTGGAGCTCTTCTGACCAACACCTCTGCCGAGGAGTGTCCAACGGGCACCCATGTCAACAT
CCTCTTCTCTCAAGACATGTGGTACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTTCAAGTTTGTGGGCAAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTGGCCACCGGCAATGCGTGTGGGGCAGGAGGAGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCACCGCCCTCTAAGAACATCTGCCAACAGC
TGGGTTACAGACTTCACACTGTGAGTTCAGACTCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTCACAGCACTGCTGAACAATGTGGCCTGGGTGGGGTTTCATCTTCTAGGGTTGAAACTAAACTGTCCA
CCCAGAAAGACACTCACCCCATTTCCCTCATTTCTTCTACACTTAAATACCTCGTGTATGGTGAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCCAAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA
TTATAGGGAATTTGGAAGTGATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFYDIEDCHGSCSDTSECTCAPGTVLGPDROTCDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDEVEGCHNNNGGCSHSLGSEKGYQCECPRLVLSNHTCQVFLCKSNAIEVNIPRELVGG
LELFLTNTSCRGVSNGTHVNILFSLKTCGTVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVFNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLSVESCFATPTSKIDEVLKYILRDGCVSDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRGAGGEDSAGLQGQTLTGGPPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCCACTGCAGGTGCTGGACTCGGCCAC
AGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCCTG
TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGCCCTCTGTGGATTCTTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCAGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
CTTCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGTGCAGTTCCTCACTACTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
CTGCGGGGGCCCCGAGACCCCAGGAGTATACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTTAATGCTGCTGCCCCTTTGAGTGTCTGGGAGCGCTTCCTTCCCTGCCCTGCCACCT
GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTTGGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCCTCGCAGCCCAGAGCGCCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGTCTCCAGCATCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAGCC
CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVIIDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAEETACRQMGSRAVEIGPDQDLVDVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNIDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCQGDSSGGPLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAACCTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC
TGTTTTTTGTTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCCTGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGAGGTAGCTGAGCCTCTTGCTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCAATTCTCCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTGTGCCAAATTATGGGTCAGAAAAGATG
GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTAGTGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAATCCA
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
TGTTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAAGTGAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHIILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKKAYDME
HTFYSGEKKKIYMEIDPVRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGCGGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTCAATGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTCCCAACATGCTGGTGACTAATTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTTACAGACAGGTACACATTTGGTGCGGCTCTGTTCTGGGCTGGGTGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCACTACAAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCTTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGACTATGTGTAAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCAAAAACAAGGAGATCCCATCTAGATTCTTCTTCTGCTTTTGAATCAGCTGGAAGTTAGAAAAGCCT
CGATTTTCTCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACTTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAAGATATTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAAGTTACCAACCAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCTCTTTCTGT
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTAAGTCCTAAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA
TGGTGGCATACACCTGATGCTCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCCTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCTAGGAAGTAGGTTAAACTAATTCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGAGTCC
AGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCCTAGAAATCGCTGGGCTGTTTCTTG
GTGGTGTGGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCCTCAGTGGAGAGTGTGGCCTTCATT
GAAACAACATCGTGGTTTTTGAACCTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAACCAAAAAAGTTATCACACCGGAAAGAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACGAGAATGAGATA
TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTATCATTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAACTTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAAATGTCCTCCA
TTTATAATGAAGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTGTCAAGGGGCTTTGCATTCAAACGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGATGGTTTTAGGAAAGTG
AAAATATTTTTGTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA
TGTCTTGTTTTTCATTGCTTACCAAAAAACAACAACAAAAAAGTTGTCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATGTCAATTTTTGTCTGTGAAAAATAAATTTCTTCTTGTGA
CCATTTCTGTTTAGTTTTACTAAATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTTATATTAATAAATTTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCGGCAGCC
GGGAGCCATGCGACCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCGTTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGGAATGTCTGAGGAAAAGCTTTGAGGAGTCTGGACACCCAACTAC
AAGCAGTGTTTCATGGAGTTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTGAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTCACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTCTTTCTCTTAGTATAGCATTTTTA
AAAAATATAAAAGCTACCAATCTTTGTACAATTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAAGVPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGACAGCTGACAGGGGCTGTGCAACTG
GCCCTAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
CTTTCTTCTGCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAAT
TACGGAACCTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
TCTTTGCTGGCATGATCAGACTCAAAGAACCTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
CTGGCCCTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGT
CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTATCAGGCAATGAGA
TCGAAGCTTTTCAAGGTTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTGTAAACTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTTGATCTGGCCAG
GGCTCTCCCAAAGCCGACGTTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCCGACGGTGGGAGCCACAGAGCCCCGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
CTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA
GATTATAAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTTGTGATAAAAAGAGCTCTTAAAGCT
GGGAAATAAGTGGTGTCTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCTTGTCCGTTTTAGTGCATTCAATAACT
GGTCATTTTCTCTCATAATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
GAACTCCGGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATT
GTTTTAAGATAAACTTCTTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTLSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAGCLG
LSLRYSNLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLSSLHLRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSTWSSLRQLDL
SGNEIEAFSGPSVFCVFNLRNLDSNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAHEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAATGACCTAGTGAAGGTTGAGATGCAAATGGAAGGAAAAAGGAACTGGAAGG
AAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCAC TTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTPVKHYLVNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRIMNP
RDKQGRGLLYKSSTDCLIQAVQGEFGMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGCCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCTCGGGGGGCATCGGCGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGACGGCTACCCCGGACTTTGATCCCCACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAAATCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACCGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCTTCATGGCTTGCCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTTCCTGTCCACCCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
TCATCTTGTCAAATGCTTCAGTTGTAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGGCTTCTTGTCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCCTGGCCAGTGGATTTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAA
AAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAARALVQQGLKVVGCAITVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSQVDICINNAGLARPDITLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVPLPSVTHFYATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC TTCTAC ATGGG CCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCTCCTTTGACTGCGGGCCGTT CAGGTGCAGAGTCTCAGTTGCCCCGGGAGC
ACCTCCCCCTCCCAGGCAGTCTGCTCAGAGGGCCTCGGCCCCAGAATTCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAAC TTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTGCTTCTGTGTCTGCGCTGTTCA TTTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAAT TGAGG AACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAAC TCTGGCCTCAAACAATCCACTTGCCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCCCTGGCCTAAACATTTATCTTTTCTTTGTGTGGGAACTTTGAAATTAT
ACAATGAATTATTGT TAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTCTGTTCTTAATTTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KKGHTLGESEMPFVKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATATAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCTGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCCTTCAGAT
TTGAAAAACTTGCTGAAAGTAGAAAGAATATACCTTATACCACAACAGTTTAGATGAATTTCCCTACCAACCT
CCCAAAGTATGTAAGAGGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTTTCCCTGTCCCGTAATCACCTTAGCACAATTCCCTGGGGTTT
GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG
GTCTCACTAGTCTAAACGCCTGGTTCTAGATGGAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
AGGTTTCGTGGGATGGCTATTAAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTTCAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATCTCTTGGAACCTTGCTCTACCTATG
ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTGATCTATAACAGAAACAATTGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTC
CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCACCCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAAACTTCTTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAACACACCATATTTCTCCTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLEFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRITITYDSLSKIPYL
EELHLLDDNSVSAVSIEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRLQYRLDMSNNNLSNLPQGIFDDLNDITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCMPMETSNNLYLFDETPVCIETETAPLRMYNPPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALALLVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKNNHSESSNRSYRDSGIPDSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCAGGCTTCTTGGA
GCCCTGCCGGGCCACTTGTCTTCAATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCGG
GTTTTGGAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGG
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAACATCATGTTCTG
TGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAGC
AATTCTCTTGCCTCATCCTCCCAGTAGCTGGGACTACAGGAGCGTGCCACCATACCTGGCTAAT
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGRRACRGSLSGLARAQGAERTVSSEQRPAMASLGLLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLSHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGFLQOSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCMGGFSDFYKLRWLEAILSQKQEGCFG
EPDAEDEELSKAIQYQQHFSSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTAGCACATGTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCAGTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACACTATACATTATGTATATTAATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVS PAYRFDPVRVDITSKGMRARYVNYIKTSE
VVR LPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMMLVLP LLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
GTTTTGACATTTCCCCGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTTTACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTCAAGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGGAGGAAAGTGATGCTG
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCAAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGCAATGATTGCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTGTCTATTTA
ATGTATTTATTTTTTACTTGGACATGAAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAATTCAGATTATATTTATAACCTG
CTAGGGGGTTATTCAATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAAATTGTGTATCTTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAA
AAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNGATTTCATNSHSDSELRPEIF
SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKDLRLSHAMTCHCGEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCGAACCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAGAAAGATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
TTCCCAAAGCATTTTGTGGAATATGATGGAACACCTCATTTCGTGTGATGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAGAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTTGGCCACATTTAGGGCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCTGAAGGAGAGAGA
GACGCCCTTACCCAAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCCCTCTCCACCGAGGAGGTGAATACACCTCTGATTTTAAATCAGTTCTGCGTTTGAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCTCCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTCCGTCATATGAAAAGATCCTGGGTTTGAACAGTTTAAATGTCAGAAAGATTGCATGG
GAACCTGGATCAGACTGTACTTGGAGGAAAAGCATTCAGAAGTCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTGCTCCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCTCTCTCTGGGTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTCGGCCGTTACGCT
CATTGTGTGCACCTCAGCTGAAAGTTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT
GCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCTGGAC
AGGGGGCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCCTCTGGGTGTCCGACACAGGTG
TTCACATCTGTGCTGTGAGGTCAGATGCCTCAGTTCTTGGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCCACCACTCATCTGCCACCCCAAGATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGCGGACGTTTCTGTGGAATTCTTAGTCTTTGGCC
TCGGACACCTTCATTGCTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCCTGCACAGCCCTCATCCCCTCTTGGCTTGAGCCGTCAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTG
CCACGCTTGACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAACT
TCCCTATAGAAAACAACTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACTACATACGCCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAAC TAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVDPKNYIRPEI
FFALASSLSSASA EKGSPI LLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLA AQESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCCTTCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGGCGCTGTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACCTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTTACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTTTGTATCCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTTTAAGAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLGEGH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCCAGCACTGCCCAGATCA
CTGAGGCCCAGGTGGTGTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGCTAACCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSWWLATVCMLLFSLHSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLLCLLALIWMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTTTGAAATTT
CAACTTTCAGATTAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLELLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
 TCCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
 CGGCCCTCTCCCGCCCCACCCACCCTCCTGGCTCTTCTGTTTTACTCCTCCTTTTCATTCTATA
 ACAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGAAGAATGGGGTT
 CCTCGGGACCGGCACTTGGATTCTGGTGTTAGTGCTCCCGATTCAAGCTTTCCCAAACCTGGAGGAA
 GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
 GCAGAAGAAGACAAGATTAATAAACATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTT
 TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
 TAAGAAGCTCCCCTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACG
 ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
 TCAACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
 AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC
 CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
 TTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAATAATACCAGAGAAAG
 TGACTCCAATGGCAGCAATTAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
 TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAACCTACAGTGAAGACAACCTTTGAGGAACTCCA
 ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTTCAGAAAAAGAGCAAAAGAGAAAGAAA
 CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT
 CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAACAAGCT
 AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAA
 CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCACAAAA
 GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
 CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGAAGATTATGACCTTT
 CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
 GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAGATCCAGGAGTCTTTCAA
 CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATCTGTGATTAAAAATTTTTGACCCAAGG
 GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
 GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFIYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISP EEGV
SYLENLDEMIALQTKNLEKNATDNISKLFPAPEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEA IKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCTCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTATCCTGGGTGTCCAGGGTGAAG
CCAGTGCTGTCTGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGTCTGGGTGAGGGGT
GAGTGGAGGAGACCATGGCGGACATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCCTTTGTCTTCTGGTTCCTCAGTTTGGATAAATCTGAGATTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGTCTACTGCTGTGGAACCTTGTAACCCTGTGGGTGAAA
CTGGGAATAACATGAAAAGATTTCTGTGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACTGACAAGTGTACCTTGAGCCCCGAGGCCAACCCTCCCCAGTTGAGCCTTATA
GGGTGAGTGTCTCCACATGAAGTCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCATA
TGCTACCTTTCCTATCTCTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTGTGAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA
GATACAATCAAAATCCAGATGCTGGTCTCTATTCCCATGAAAAGTGTCTATGACATATTGAGA
AGACCTACTTACAAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATAACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTCAGGGTGGTGGCAGTAT
AGGTGATTTTCTTTTAAATCTGTGTAATTTATCTGTATTTCTTAATTTTCTACAATGAAGATGA
ATTCTTTGTATAAAAAAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAAGAACTTAAAGAACAAAATCATCTGGTAATTCTTTCTAGAA
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAACTCTCAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTATGCTGGATGAAGGTAGACCTAAATTCATAT
GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAGAGGCAAAGAAGAAATCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAAATAAATTCGGCTGTTTTAA
GCCACCAAGGATAAATGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGTCCCCCTCCCAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATAAGGAGAGTATGATGGCATGCCACTTCCAAGATTAGGTATATAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTGCCCACCCACCGCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCTTATAAAGAGACTTACGTGGTAAAAATGAAGTCTCCT
GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSTTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGWGNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGCGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAAGTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTATTGTAAGTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQSVSSFLMGTLATSCLLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHYSHWPSCCPSKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTCAATGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAACCTTCAACACCTCCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTGCAGAAGTGAGGTCATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLRDLRVEPVTTSVATGDYSILMNVSWV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVEINTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACCTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16o

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMsrNIESRSTSPWNYTWTWDPNRYPSVVQAQCRNLGCINAQKEDISMNSVPIQQETLVV
RRKHQGCsvSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTTGAGGGCTGAAGAGGGATTTC
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTCTTGCTGTCTT
TGGCACTGGGGCCGAAGCCAGTGGTCTTTCTCTGGAGAGGCTTGTGGGGCCTCAGGACGCTACC
CACTGCTCTCCGGGCTCTCCTGCCGCTCTGGGACAGTGACATACTCTGCCCTGCCCTGGGGACAT
CGTGCCCTGCTCCGGGGCCCGTGTCTGGCGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCTCTCTCCAGGCCCAAGTCGTGCTCTCCTTCCAGGCCTACCTACTGCCCGCTGCGTCTGCTG
TGGAGGTGCAAGTGCCCTGCTGCCCTTGTGCACTTGGTTCAGTCTGTGGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGGGTACGAATCTGGTCTTACTCAGCCCAGGTACGAGAA
GGAATCAACACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTGACGAGATGGTGACA
ACGTGCACTGTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCCAAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTTCCTGCTCTGTATTTCAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCCCTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCCCGACTG
CGACTGCTGACCTGCAGAGCTGGCTGCTGGACGCACCGTGTCTGCTGCCCGCAGAAGCGGCACT
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCTGGTCCCACCGCTTTCCTGGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTTCCCATTGCTGAAAGGCCACCTAACCTCTGTGTTTCAGGTG
AACAGCTCGGAGAAGCTGCAGCTGCAGGAGTGTCTGTGGGCTGACTCCCTGGGGCCTCTCAAAGA
CGATGTGCTACTGTTGGAGACACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCTTGGAAACCA
GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTG
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCTACTCTTTGCCG
CTGCGCTTTCCCTCATCTCTCTTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGA
CAGGACGTCCGCTCGGGGGCGGCCGCCAGGGGCCGCGCGGCTCTGCTCCTCTACTCAGCCGATGA
CTCGGGTTTCGAGCGCTTGGTGGGGCGCCCTGGCGTGGGCCCTGTGCCAGCTGCCGCTGCGCGTGG
CCGTAGACCTGTGGAGCCGTCTGTAACTGAGCGCGCAGGGGCCGTGGCTTGGTTTCACGCGCAG
CGGCGCCAGACCTGCAGGAGGGCGCGGTGGTGGTCTTGTCTCTCTCCCGGTGCGGTGGCGCT
GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGGCGCACGGCCCGCACGACGCTTCC
GCGCCTCGCTCAGCTGCGTGTGCCGACTTCTTGACAGGGCCGGGCGCCCGGACGTACGTGGGG
GCCTGCTTCGACAGGCTGTCTCCACCGGACGCCGTACCCGCCCTTTTCCGCACCGTGCCCGTCTT
CACACTGCCCTCCCACTGCCAGACTTCTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCGTTCCG
GGCGGCTCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGGATAGCTACTTC
CATCCCCCGGGGACTCCCGCGCGGACGCGGGTGGGACCAGGGGCGGGACCTGGGGCGGGGA
CGGGACTTAAATAAAGGCAGACGCTGTTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELRLVGPQDATHCSPGLSCRLWSDILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSTQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVNSADGDNVHLVLNV
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQ
LWQAARLRLLTLQSWLLDAPCSLPAAEALCWRAFGGDPCLVPLSWENVTVDKVLEFPLKGGHNPCLVQ
VNSSEKLQLECLWADSLGPLKDDVLLLETRGPDNRSLCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSILLKDHAKGWLRLKQDVRSGAAARG
RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDLFGALQQPRAPRSRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCCTGC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
 GTCAGTGGGGAGGCGGTCAGCCACCAAGATGACTGACAGGTCAGCTCTCTGCAGCACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTAGATGATTGTTTCATCCTA
 CCCCCACGCCAATCCGTGCAGGCGATGGGCCACCGCTAACCCCTGGAAGACATCTTCCATGACCTG
 TTCTACCACTTAGAGCTCAGGTCAACCGCACCTACCAAATGCACCTTGAGGGAAGCAGAGAGA
 ATATGAGTTCTTCGGCTGACCCCTGACACAGAGTTCTTGCCACCATCATGATTTGCGTTCCCA
 CCTGGGCCAAGGAGAGTCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTCTCCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
 CTACAGATATGTACCAAGCCGCTGCACCTCCCACTCCCTGAACGTCCAGCGAGTCTGACTT
 TCCAGCCGCTGCGCTTCATCCAGGACGACGTCCCTGATCCCTGTCTTGACCTCAGCGGCCCGAGC
 AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
 GTCGGGCCCCCATCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCCATTCTACGCCCCACA
 GGCCATCTTAAGGTCAGCCTTCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGCATGGAAGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTTAA
 CACCTTAGGCCTAAAGGTCAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGCAGGAGGTGACCTCCTTGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC
 CCCTGGGGATTTCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAACTGGATTCTCTTTTCAGAGGCTGGCCCTGACTGTGCAGTGGGAGTCTG
 AGGGGAATGGGAAAGGCTTGGTGCTTCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
 AGAGGGAGTGGCATGCAGGGCCCCCTGCCATGGGTGCGCTCCTACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGTCTGAGCCCTG
 CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCGAGGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCTCTCCTTGCCCCATTCTGGCCAGTTTC
 ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCCTTCTGTCTATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACCAGGCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCTGCAGCTCA
 TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTAGCTTCATTCTCTGATAGAACAAGC
 GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTCTGCAGGCAGGAGTTTCAGACCT
 ATCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
 TGTACTGATGTACAACCTTGAAGCTCTGCCTTGGGTTTCAGCCCATCTGGGCTCAAATTCAGC
 CTCACCACTACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCTC
 ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCTTGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTelyarvtavsaggrsatkmtDRFSSLQHTTLKPPDVCIS
KVRSIQMIHVHTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCrvKTLpDRTWTYsfsgaflfsmgflvavlcylsyrYVTKPPAP
PNSLNVQrVLTfQPLRFIqEHVLIpVFDLSGPSSLAQPvQYSQIRVSGPREPAGAPQRHSLSEIT
YLQPDISILQPSNVPPPQILSPLSYAPNAapevgPPSYAPQVTPEAQFFYAPQAIskVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRtSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHFMSLPLQPPSGPC
SPSDQGSPWGLLESVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTvQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCGCGGTGGCCACAACATGG
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTCTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTGGGCTGGA
AGTGTGTAACACAGTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCTGTGCCTGAGCCCGAGGCATTTCAGAGCTGATTTCAGAGGATGGAGAAGGTGCTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTTTCGTTGGCATCAAATCTAAGTTTGTGTTTACAAAGATTGTTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPD CRFVN
FKKGDDVYVYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPAETDFVCFEGGRD
DFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRA DSEGE GA
FSESTEG LQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCCTAACTTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCCGCCGCGTGGTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAACT
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCCTCCTCA
GCCTTGAGAAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAACATCTGG
TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATCCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTTACCAGTGAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAACTAGTCCAAACAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAAGTAATTTTTTATATGTCTATTATTTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAACCTTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAAGCAATTCCTTTTATTTCTTTCACCTTTCATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGGTTGATAAAA
ATGAACTGTTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
AATAAGAAGCTATTTTCATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT
TTGTAATAATGTAACCTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAAQAAAAGAAG
TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCTGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTGATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAAATGGGTTGG
AACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGGCTTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTTCTTCTCAAGAGAAAGTTGTAACCTCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAATAAAGAGTTCTTGTCTTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRRVVSQGKVCFADEFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKS KGRTKTS PNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217